

RECEIVED
APR 18 2002
TECH CENTER 1600/2900



OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/090,624

DATE: 03/19/2002

TIME: 14:47:20

Input Set : N:\Cr3\RULE60\10090624.raw

Output Set: N:\CRF3\03192002\J090624.raw

RECEIVED

APR 18 2002

TECH CENTER 1600/2900

1 <110> APPLICANT: TAKAKURA, Hikaru
2 MORISHITA, Mio
3 SHIMOJO, Tomoko
4 ASADA, Kiyozo
5 KATO, Ikunoshin
6 <120> TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
7 <130> FILE REFERENCE: TAKAKURA=6
8 <140> CURRENT APPLICATION NUMBER: 10/090,624
9 <141> CURRENT FILING DATE: 2002-03-06
10 <150> PRIOR APPLICATION NUMBER: 09/445,472
11 <151> PRIOR FILING DATE: 1999-12-06
12 <150> PRIOR APPLICATION NUMBER: 151969/1997
13 <151> PRIOR FILING DATE: 1997-06-10
14 <160> NUMBER OF SEQ ID NOS: 33
15 <170> SOFTWARE: PatentIn version 3.0
17 <210> SEQ ID NO: 1
18 <211> LENGTH: 412
19 <212> TYPE: PRT
20 <213> ORGANISM: Pyrococcus furiosus
21 <400> SEQUENCE: 1

ENTERED

22	Ala	Glu	Leu	Glu	Gly	Leu	Asp	Glu	Ser	Ala	Ala	Gln	Val	Met	Ala	Thr
23	1			5					10					15		
24	Tyr	Val	Trp	Asn	Leu	Gly	Tyr	Asp	Gly	Ser	Gly	Ile	Thr	Ile	Gly	Ile
25			20					25					30			
26	Ile	Asp	Thr	Gly	Ile	Asp	Ala	Ser	His	Pro	Asp	Leu	Gln	Gly	Lys	Val
27			35				40						45			
28	Ile	Gly	Trp	Val	Asp	Phe	Val	Asn	Gly	Arg	Ser	Tyr	Pro	Tyr	Asp	Asp
29		50					55				60					
30	His	Gly	His	Gly	Thr	His	Val	Ala	Ser	Ile	Ala	Ala	Gly	Thr	Gly	Ala
31	65					70				75				80		
32	Ala	Ser	Asn	Gly	Lys	Tyr	Lys	Gly	Met	Ala	Pro	Gly	Ala	Lys	Leu	Ala
33				85					90					95		
34	Gly	Ile	Lys	Val	Leu	Gly	Ala	Asp	Gly	Ser	Gly	Ser	Ile	Ser	Thr	Ile
35			100					105					110			
36	Ile	Lys	Gly	Val	Glu	Trp	Ala	Val	Asp	Asn	Lys	Asp	Lys	Tyr	Gly	Ile
37		115					120					125				
38	Lys	Val	Ile	Asn	Leu	Ser	Leu	Gly	Ser	Ser	Gln	Ser	Ser	Asp	Gly	Thr
39		130					135					140				
40	Asp	Ala	Leu	Ser	Gln	Ala	Val	Asn	Ala	Ala	Trp	Asp	Ala	Gly	Leu	Val
41	145				150					155				160		
42	Val	Val	Val	Ala	Ala	Gly	Asn	Ser	Gly	Pro	Asn	Lys	Tyr	Thr	Ile	Gly
43				165				170						175		
44	Ser	Pro	Ala	Ala	Ala	Ser	Lys	Val	Ile	Thr	Val	Gly	Ala	Val	Asp	Lys

RAW SEQUENCE LISTING

DATE: 03/19/2002

PATENT APPLICATION: US/10/090,624

TIME: 14:47:20

Input Set : N:\Crif3\RULE60\10090624.raw

Output Set: N:\CRF3\03192002\J090624.raw

```

45          180          185          190
46 Tyr Asp Val Ile Thr Ser Phe Ser Ser Arg Gly Pro Thr Ala Asp Gly
47          195          200          205
48 Arg Leu Lys Pro Glu Val Val Ala Pro Gly Asn Trp Ile Ile Ala Ala
49          210          215          220
50 Arg Ala Ser Gly Thr Ser Met Gly Gln Pro Ile Asn Asp Tyr Tyr Thr
51          225          230          235          240
52 Ala Ala Pro Gly Thr Ser Met Ala Thr Pro His Val Ala Gly Ile Ala
53          245          250          255
54 Ala Leu Leu Leu Gln Ala His Pro Ser Trp Thr Pro Asp Lys Val Lys
55          260          265          270
56 Thr Ala Leu Ile Glu Thr Ala Asp Ile Val Lys Pro Asp Glu Ile Ala
57          275          280          285
58 Asp Ile Ala Tyr Gly Ala Gly Arg Val Asn Ala Tyr Lys Ala Ile Asn
59          290          295          300
60 Tyr Asp Asn Tyr Ala Lys Leu Val Phe Thr Gly Tyr Val Ala Asn Lys
61          305          310          315          320
62 Gly Ser Gln Thr His Gln Phe Val Ile Ser Gly Ala Ser Phe Val Thr
63          325          330          335
64 Ala Thr Leu Tyr Trp Asp Asn Ala Asn Ser Asp Leu Asp Leu Tyr Leu
65          340          345          350
66 Tyr Asp Pro Asn Gly Asn Gln Val Asp Tyr Ser Tyr Thr Ala Tyr Tyr
67          355          360          365
68 Gly Phe Glu Lys Val Gly Tyr Tyr Asn Pro Thr Asp Gly Thr Trp Thr
69          370          375          380
70 Ile Lys Val Val Ser Tyr Ser Gly Ser Ala Asn Tyr Gln Val Asp Val
71          385          390          395          400
72 Val Ser Asp Gly Ser Leu Ser Gln Pro Gly Ser Ser
73          405          410

```

75 <210> SEQ ID NO: 2

76 <211> LENGTH: 1236

77 <212> TYPE: DNA

78 <213> ORGANISM: Artificial Sequence

79 <220> FEATURE:

80 <223> OTHER INFORMATION: Synthetic

81 <400> SEQUENCE: 2

```

82 gcagaattag aaggactgga tgagtctgca gctcaagtta tggcaactta cgtttggaac 60
83 ttgggatatg atggttctgg aatcacaata ggaataattg aacttggaat tgacgcttct 120
84 catccagatc tccaaggaaa agtaattggg tgggtagatt ttgtcaatgg taggagttat 180
85 ccatacgaatg accatggaca tgggaactcat gtagcttcaa tagcagctgg tactggagca 240
86 gcaagtaatg gcaagtacaa ggggaatggc ccaggagcta agctggcggg aattaagggtt 300
87 ctaggtgccg atggttctgg aagcatatct actataatta agggagtga gtgggccgtt 360
88 gataacaaag ataagtacgg aattaaggctt attaatcttt ctcttggttc aagccagagc 420
89 tcagatggta ctgacgctct aagtcaggct gttaatgcag cgtgggatgc tggattagtt 480
90 gttgtggttg ccgctggaaa cagtggacct aacaagtata caatcggttc tccagcagct 540
91 gcaagcaaag ttattacagt tggagccgtt gacaagtagt atgttataac aagcttctca 600
92 agcagagggc caactgcaga cggcaggctt aagcctgagg ttgttgctcc aggaaactgg 660
93 ataattgctg ccagagcaag tgggaactagc atgggtcaac caattaatga ctattacaca 720
94 gcagctcctg ggacatcaat ggcaactcct cacgtagctg gtattgcagc cctcttgctc 780

```

RAW SEQUENCE LISTING

DATE: 03/19/2002

PATENT APPLICATION: US/10/090,624

TIME: 14:47:20

Input Set : N:\Crif3\RULE60\10090624.raw

Output Set: N:\CRF3\03192002\J090624.raw

```
95      caagcacacc cgagctggac tccagacaaa gtaaaaacag ccctcataga aactgctgat      840
96      atcgtaaagc cagatgaaat agccgatata gcctacgggtg caggtaggggt taatgcatac      900
97      aaggctataa actacgataa ctatgcaaag ctagtgttca ctggatatgt tgccaacaaa      960
98      ggcagccaaa ctcaccagtt cgttattagc ggagcttcgt tcgtaactgc cacattatac      1020
99      tgggacaatg ccaatagcga ccttgatctt tacctctacg atoccaatgg aaaccaggtt      1080
100     gactactctt acaccgccta ctatggattc gaaaagggtt gttattacaa cccaactgat      1140
101     ggaacatgga caattaaggt tgtaagctac agcggaagtg caaactatca agtagatgtg      1200
102     gtaagtgatg gttccctttc acagcctgga agttca      1236
104 <210> SEQ ID NO: 3
105 <211> LENGTH: 29
106 <212> TYPE: PRT
107 <213> ORGANISM: Bacillus subtilis
108 <400> SEQUENCE: 3
109     Met Arg Ser Lys Lys Leu Trp Ile Ser Leu Leu Phe Ala Leu Thr Leu
110     1          5          10          15
111     Ile Phe Thr Met Ala Phe Ser Asn Met Ser Ala Gln Ala
112     20          25
114 <210> SEQ ID NO: 4
115 <211> LENGTH: 522
116 <212> TYPE: PRT
117 <213> ORGANISM: Pyrococcus furiosus
118 <220> FEATURE:
119 <221> NAME/KEY: misc_feature
120 <222> LOCATION: (428)..(428)
121 <223> OTHER INFORMATION: Xaa at position 428 is Gly or Val.
122 <400> SEQUENCE: 4
123     Ala Glu Leu Glu Gly Leu Asp Glu Ser Ala Ala Gln Val Met Ala Thr
124     1          5          10          15
125     Tyr Val Trp Asn Leu Gly Tyr Asp Gly Ser Gly Ile Thr Ile Gly Ile
126     20          25          30
127     Ile Asp Thr Gly Ile Asp Ala Ser His Pro Asp Leu Gln Gly Lys Val
128     35          40          45
129     Ile Gly Trp Val Asp Phe Val Asn Gly Arg Ser Tyr Pro Tyr Asp Asp
130     50          55          60
131     His Gly His Gly Thr His Val Ala Ser Ile Ala Ala Gly Thr Gly Ala
132     65          70          75          80
133     Ala Ser Asn Gly Lys Tyr Lys Gly Met Ala Pro Gly Ala Lys Leu Ala
134     85          90          95
135     Gly Ile Lys Val Leu Gly Ala Asp Gly Ser Gly Ser Ile Ser Thr Ile
136     100         105         110
137     Ile Lys Gly Val Glu Trp Ala Val Asp Asn Lys Asp Lys Tyr Gly Ile
138     115         120         125
139     Lys Val Ile Asn Leu Ser Leu Gly Ser Ser Gln Ser Ser Asp Gly Thr
140     130         135         140
141     Asp Ala Leu Ser Gln Ala Val Asn Ala Ala Trp Asp Ala Gly Leu Val
142     145         150         155         160
143     Val Val Val Ala Ala Gly Asn Ser Gly Pro Asn Lys Tyr Thr Ile Gly
144     165         170         175
145     Ser Pro Ala Ala Ala Ser Lys Val Ile Thr Val Gly Ala Val Asp Lys
```

RAW SEQUENCE LISTING

DATE: 03/19/2002

PATENT APPLICATION: US/10/090,624

TIME: 14:47:20

Input Set : N:\Crf3\RULE60\10090624.raw

Output Set: N:\CRF3\03192002\J090624.raw

```
146          180          185          190
147      Tyr Asp Val Ile Thr Ser Phe Ser Ser Arg Gly Pro Thr Ala Asp Gly
148          195          200          205
149      Arg Leu Lys Pro Glu Val Val Ala Pro Gly Asn Trp Ile Ile Ala Ala
150          210          215          220
151      Arg Ala Ser Gly Thr Ser Met Gly Gln Pro Ile Asn Asp Tyr Tyr Thr
152      225          230          235          240
153      Ala Ala Pro Gly Thr Ser Met Ala Thr Pro His Val Ala Gly Ile Ala
154          245          250          255
155      Ala Leu Leu Leu Glu Ala His Pro Ser Trp Thr Pro Asp Lys Val Lys
156          260          265          270
157      Thr Ala Leu Ile Glu Thr Ala Asp Ile Val Lys Pro Asp Glu Ile Ala
158          275          280          285
159      Asp Ile Ala Tyr Gly Ala Gly Arg Val Asn Ala Tyr Lys Ala Ile Asn
160          290          295          300
161      Tyr Asp Asn Tyr Ala Lys Leu Val Phe Thr Gly Tyr Val Ala Asn Lys
162      305          310          315          320
163      Gly Ser Gln Thr His Gln Phe Val Ile Ser Gly Ala Ser Phe Val Thr
164          325          330          335
165      Ala Thr Leu Tyr Trp Asp Asn Ala Asn Ser Asp Leu Asp Leu Tyr Leu
166          340          345          350
167      Tyr Asp Pro Asn Gly Asn Gln Val Asp Tyr Ser Tyr Thr Ala Tyr Tyr
168          355          360          365
169      Gly Phe Glu Lys Val Gly Tyr Tyr Asn Pro Thr Asp Gly Thr Trp Thr
170          370          375          380
171      Ile Lys Val Val Ser Tyr Ser Gly Ser Ala Asn Tyr Gln Val Asp Val
172      385          390          395          400
173      Val Ser Asp Gly Ser Leu Ser Gln Pro Gly Ser Ser Pro Ser Pro Gln
174          405          410          415
W--> 175      Pro Glu Pro Thr Val Asp Ala Lys Thr Phe Gln Xaa Ser Asp His Tyr
176          420          425          430
177      Tyr Tyr Asp Arg Ser Asp Thr Phe Thr Met Thr Val Asn Ser Gly Ala
178          435          440          445
179      Thr Lys Ile Thr Gly Asp Leu Val Phe Asp Thr Ser Tyr His Asp Leu
180          450          455          460
181      Asp Leu Tyr Leu Tyr Asp Pro Asn Gln Lys Leu Val Asp Arg Ser Glu
182      465          470          475          480
183      Ser Pro Asn Ser Tyr Glu His Val Glu Tyr Leu Thr Pro Ala Pro Gly
184          485          490          495
185      Thr Trp Tyr Phe Leu Val Tyr Ala Tyr Tyr Thr Tyr Gly Trp Ala Tyr
186          500          505          510
187      Tyr Glu Leu Thr Ala Lys Val Tyr Tyr Gly
188          515          520
190 <210> SEQ ID NO: 5
191 <211> LENGTH: 4765
192 <212> TYPE: DNA
193 <213> ORGANISM: Artificial Sequence
194 <220> FEATURE:
195 <223> OTHER INFORMATION: Synthetic
```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/090,624

DATE: 03/19/2002

TIME: 14:47:20

Input Set : N:\Crif3\RULE60\10090624.raw

Output Set: N:\CRF3\03192002\J090624.raw

196 <400> SEQUENCE: 5

197	tttaaattat	aagatataat	cactccgagt	gatgagtaag	atacatcatt	acagtcccaa	60
198	aatgtttata	attggaacgc	agtgaatata	caaaatgaat	ataacctcgg	aggtgactgt	120
199	agaatgaata	agaagggact	tactgtgcta	tttatagcga	taatgctcct	ttcagtagtt	180
200	ccagtgcact	ttgtgtccgc	agaaacacca	ccggttagtt	cagaaaattc	aacaacttct	240
201	atactcccta	accaacaagt	tgtgacaaaa	gaagtttcac	aagcggcgct	taatgctata	300
202	atgaaaggac	aacccaacat	ggttcttata	atcaagacta	aggaaggcaa	acttgaagag	360
203	gcaaaaaccg	agcttgaaaa	gctaggtgca	gagattcttg	acgaaaatag	agttcttaac	420
204	atgttgctag	ttaagattaa	gcctgagaaa	gttaaagagc	tcaactatat	ctcatctctt	480
205	gaaaaagcct	ggcttaacag	agaagttaag	ctttcccctc	caattgtcga	aaaggacgtc	540
206	aagactaagg	agccctccct	agaacaaaa	atgtataaca	gcacctgggt	aattaatgct	600
207	ctccagttca	tccaggaatt	tggatatgat	ggtagtgggt	ttgttgttgc	agtacttgac	660
208	acgggagttg	atccgaacca	tcctttcttg	agcataactc	cagatggacg	caggaaaatt	720
209	atagaatgga	aggattttac	agacgagggg	ttcgtggata	catcattcag	ctttagcaag	780
210	gttgtaaatg	ggactcttat	aattaacaca	acattccaag	tggcctcagg	tctcacgctg	840
211	aatgaatcga	caggacttat	ggaatacgtt	gttaagactg	tttacgtgag	caatgtgacc	900
212	attggaata	tcacttctgc	taatggcatc	tatcacttcg	gcctgctccc	agaaagatac	960
213	ttcgacttaa	acttcgatgg	tgatcaagag	gacttcctatc	ctgtcttatt	agttaactcc	1020
214	actggcaatg	gttatgacat	tgcataatgt	gatactgacc	ttgactacga	cttcaccgac	1080
215	gaagttccac	ttggccagta	caacgttact	tatgatgttg	ctgttttttag	ctactactac	1140
216	ggctctctca	actacgtgct	tgcagaaata	gatcctaacg	gagaatatgc	agtatttggg	1200
217	tgggatggtc	acggtcacgg	aactcacgta	gctggaactg	ttgctgggta	cgacagcaac	1260
218	aatgatgctt	gggattggct	cagtatgtac	tctggtgaat	gggaagtgtt	ctcaagactc	1320
219	tatggttggg	attatacgaa	cgttaccaca	gacaccgtgc	aggggtgttg	tccaggtgcc	1380
220	caaataatgg	caataagagt	tcttaggagt	gatggacggg	gtagcatgtg	ggatattata	1440
221	gaaggatga	catacgagc	aacccatggt	gcagacgtta	taagcatgag	tctcggtgga	1500
222	aatgctccat	acttagatgg	tactgatcca	gaaagcgttg	ctgtggatga	gcttaccgaa	1560
223	aagtaoagg	ttgtattcgt	aatagctgca	ggaaatgaag	gtcctggcat	taacatcggt	1620
224	ggaagtccct	gtgttgcaac	aaaggcaata	actgttggag	ctgctgcagt	gcccattaac	1680
225	gttggagttt	atgtttccca	agcacttgga	tatcctgatt	actatggatt	ctattacttc	1740
226	cccgcctaca	caaacgttag	aatagcattc	ttctcaagca	gagggccgag	aatagatggt	1800
227	gaaataaaa	ccaatgtagt	ggctccaggt	tacgggaattt	actcatccct	gccgatgtgg	1860
228	attggcggag	ctgacttcat	gtctggaact	tcgatggcta	ctccacatgt	cagcgggtgtc	1920
229	gttgcaactc	tcataagcgg	ggcaaaggcc	gaggggaatat	actacaatcc	agatataatt	1980
230	aagaagggtc	ttgagagcgg	tgcaacctgg	cttgaggagg	atccatatac	tgggcagaag	2040
231	tacactgagc	ttgaccaagg	tcatggtctt	gttaacgtta	ccaagtcctg	ggaaatcctt	2100
232	aaggctataa	acggcaccac	tctcccaatt	gttgatcaact	gggcagacaa	gtcctacagc	2160
233	gactttgcgg	agtacttggg	tgtggacggt	ataagagggtc	tctacgcaag	gaactctata	2220
234	cctgacattg	tcgagtggca	cattaagtac	gtagggggaca	cggagtacag	aacttttgag	2280
235	atctatgcaa	ctgagccatg	gattaagcct	tttgtcagtg	gaagtgtaat	tctagagaac	2340
236	aataccgagt	ttgtccttag	ggtgaaatat	gatgtagagg	gtcttgagcc	aggtctctat	2400
237	gttggaagg	taatcattga	tgatccaaca	acgccagtta	ttgaagacga	gatcttgaac	2460
238	acaattgtta	ttcccagaga	gttcaactcct	gagaacaatt	acaccctcac	ctggtatgat	2520
239	attaatgggt	cagaaatggg	gactcaccac	ttcttcaactg	tgcttgaggg	agtggacggt	2580
240	ctctacgcga	tgaccacata	ctgggactac	ggtctgtaca	gaccagatgg	aatgtttgtg	2640
241	ttcccatacc	agctagatta	tcttcccgcct	gcagtctcaa	atccaatgcc	tggaactggg	2700
242	gagctagtat	ggactggatt	taactttgca	cccctctatg	agtcgggctt	ccttgtaagg	2760
243	atttacggag	tagagataac	tccaagcgtt	tggtacatta	acaggacata	ccttgacact	2820
244	aacactgaat	tctcaattga	attcaatatt	actaacatct	atgcccacat	taatgcaact	2880

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/090,624

DATE: 03/19/2002
TIME: 14:47:21

Input Set : N:\Crf3\RULE60\10090624.raw
Output Set: N:\CRF3\03192002\J090624.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:4; Xaa Pos. 428
Seq#:24; N Pos. 20,21,22,23,24,25
Seq#:25; N Pos. 22,23,24,25,26,27

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:32; Line(s) 970

VERIFICATION SUMMARY

DATE: 03/19/2002

PATENT APPLICATION: US/10/090,624

TIME: 14:47:21

Input Set : N:\Crf3\RULE60\10090624.raw

Output Set: N:\CRF3\03192002\J090624.raw

L:175 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:416
L:843 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:846 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:24
L:847 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:24 after pos.:0
L:855 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:858 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:25
L:859 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25 after pos.:0
L:967 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:969 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:32
L:971 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:32
L:981 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:983 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:33
L:985 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:33